J. Mol. Biol. (1995) **252**, 563–571

JMB

Formation of Stable and Functional HIV-1 Nucleoprotein Complexes *in Vitro*

Valérie Tanchou, Caroline Gabus, Véronique Rogemond and Jean-Luc Darlix*

LaboRetro, Unité de Virologie humaine (INSERM-ENS 412), Ecole Normale Supérieure, 46 allée d'Italie 69364 Lyon, France

**Corresponding author*

HIV genomic RNA resides within the nucleocapsid, in the interior of the virus, which serves to protect the RNA against nuclease degradation and to promote its reverse transcription. To investigate the role of nucleocapsid protein (NCp7) in the stability and replication of genomic RNA within the nucleocapsid, we used NCp7, reverse transcriptase (RT) and RNAs representing the 5' and 3' regions of the genome to reconstitute functional HIV-1 nucleocapsids. The nucleoprotein complexes generated *in vitro* were found to be stable, which, according to biochemical and genetic data, probably results from the tight binding of NCp7 molecules to the RNA and strong NCp7/NCp7 interactions. The nucleoprotein complexes efficiently protected viral RNA against RNase degradation and, at the same time, promoted viral DNA synthesis by RT. DNA strand transfer from the 5' to the 3' RNA template was very efficient in nucleoprotein complexes formed in the presence of both RNAs, but not when the RNAs were in separate complexes. These results indicate that the *in vitro* reconstituted HIV-1 nucleoprotein complexes function like virion nucleocapsids and thus provide a way to study at the molecular level this viral substructure and the synthesis of proviral DNA, and to search for new anti-HIV agents.

Keywords: HIV-1; nucleocapsid protein; protein–protein interactions; reverse transcription

Introduction

Retroviruses are a family of ubiquitous animal viruses 110 to 140 nm in diameter, composed of an inner core and an outer envelope formed of a lipid bilayer of cellular origin in which are anchored the surface and transmembrane viral glycoproteins. The core is formed of an outer shell of capsid protein molecules surrounding the interior nucleocapsid, which contains the dimeric RNA genome closely associated with a large number of nucleocapsid (NC) protein molecules and a few molecules of cellular tRNAs, reverse transcriptase (RT) and integrase (IN) (Chen *et al.*, 1980; Coffin, 1985; Darlix *et al.*, 1990; Darlix, 1991).

As retroviral replication most probably occurs within the nucleocapsid (Chen *et al.*, 1980; Coffin, 1985; Hu & Temin, 1990; Darlix, 1991), this nucleoprotein complex must protect the genomic RNA to ensure viral stability, replication and dissemination. Protection is, however, generally incomplete, since the virion genomic RNA often contains nicks (Coffin, 1979, 1985; Darlix *et al.*, 1990; Darlix, 1991). Conversion of the genomic single-stranded RNA into a double-stranded proviral DNA by RT is a complex process that necessitates two DNA strand transfers (Coffin, 1985; Hu & Temin, 1990) and possibly additional transfers in order to bypass nicks present in the virion genomic RNA (Coffin, 1979, 1985). HIV-1 NCp7, like Rous sarcoma virus NCp12 and murine leukemia virus NCp10, has nucleic acid binding and annealing activities that promote the annealing of the replication primer tRNA onto the primer-binding-site (PBS; Darlix *et al.*, 1990; Méric *et al.*, 1988; de Rocquigny *et al.*, 1992, 1993; Prats *et al.*, 1988; Dib-Hajj *et al.*, 1993; Barat *et al.*, 1989; Bieth *et al.*, 1990) and the first DNA strand transfer during proviral DNA synthesis (Darlix *et al.*, 1993; Allain *et al.*, 1994; Lapadat-Tapolsky *et al.*, 1993; You & Henry, 1994).

In order to study the structure–function relationships of the HIV-1 nucleocapsid, we devised an

Abbreviations used: NC, nucleocapsid; RT, reverse transcriptase; IN, integrase; PBS, primer binding site; DSP, bithiobis(succinimidylpropionate); ss-cDNA, minus strand strong stop cDNA; RNAoligo, RNA oligonucleotide; PCR, polymerase chain reaction; TAR, *trans*-acting responsive element.

0022–2836/95/400563–09 $12.00/0

in vitro system allowing the formation of functional HIV-1 nucleoprotein complexes. Viral RNAs representing the 5' and 3' regions of the genomic RNA were mixed with NCp7 protein and RT. The nature of the NCp7/RNA and NCp7/NCp7 interactions, the level of viral RNA protection against nuclease attack, and the extent of proviral DNA synthesis and DNA strand transfer in the nucleoprotein complexes, were examined.

Results

Formation of stable HIV-1 nucleoprotein complexes *in vitro*

Recently it has been reported that high molecular mass complexes can be formed by incubating HIV-1 NCp7, viral RNAs representing the 5' and 3' regions of genomic RNA, primer $tRNA^{Lys,3}$ and RT (Darlix *et al.*, 1993). Viral DNA synthesis with DNA strand transfer was found to be extensive in these nucleoprotein complexes of high molecular mass.

To examine the stability of these *in vitro* generated HIV-1 nucleoprotein complexes, gel retardation assays were performed in the presence of the denaturing agent sodium dodecyl sulfate (SDS). The nucleoprotein complexes were found to be resistant to 0.1% SDS, and most, if not all, viral RNA was present in complexes with about one NCp7 molecule per 15 nucleotides (Figure 1, lanes 5 to 7). The presence of both free RNA and nucleoprotein complexes at limiting NC protein to RNA ratios (lanes 5, 6 and 8) is in agreement with the observation that binding of NC protein to the viral RNA containing the encapsidation-dimerization sequences (E-DLS) is co-operative (Bieth *et al.*, 1990; Darlix *et al.*, 1990; Lever *et al.*, 1989; Darnull *et al.*, 1994). Gel retardation assays were also performed using the peptides NCp7-A and NCp7-B (de Rocquigny *et al.*, 1992). NCp7-A lacks the N-terminal region and the first zinc finger, and still binds to RNA, though with a reduced affinity, but has little nucleic acid annealing activity. NCp7-B lacks both zinc fingers and exhibits a high level of nucleic acid binding and annealing activities (de Rocquigny *et al.*, 1992). Clearly, NCp7-A did not form complexes resistant to 0.1% SDS (lanes 2 to 4) while NCp7-B was able to form stable nucleoprotein complexes (lanes 8 to 10). The binding of the RT enzyme alone to viral RNA did not lead to the formation of stable complexes under these conditions (data not shown). To verify that most of the NCp7 molecules were within the nucleoprotein complexes of high molecular mass, the complexes were pelleted by centrifugation (Lapadat-Tapolsky *et al.*, 1993; Bieth *et al.*, 1990), and the NCp7 content of the pellets and the supernatants was analysed by SDS-PAGE. When one molecule of NCp7 was added per 10 to 20 nucleotides of viral RNA, most of the NCp7 protein was in the pellet (data not shown).

Previous data on Rous sarcoma virus (RSV) have suggested that the overall structure of the viral nucleocapsid depends on NC protein/genomic RNA as well as NC protein/NC protein interactions (Darlix & Spahr, 1982; Méric *et al.*, 1984). To examine interactions between NCp7 molecules in HIV-1 nucleocapsids, we first used a biochemical approach. Nucleoprotein complexes were formed using ^{32}P-labelled viral RNA representing the 5' leader of HIV-1 RNA and NCp7. The complexes were irradiated with ultraviolet light at 252 nm to cross-link NCp7 molecules to the viral RNA (Darlix *et al.*, 1990, and references therein, de Rocquigny *et al.*, 1992; Méric *et al.*, 1984) and then treated with dithiobis(succinimidylpropionate; DSP, or Lomant's reagent), an agent that cross-links proteins in close contact (Schweizer *et al.*, 1982, and references therein). The viral RNA was then extensively digested with T_1 RNase and the RNAoligo/NC protein complexes analysed by SDS-PAGE and autoradiography (see Figure 2(a) for details). As previously reported (Darlix *et al.*, 1990; de Rocquigny *et al.*, 1992), ^{32}P]RNAoligo/NCp7 complexes of 25 to 35 kDa were observed (Figure 2(b), lane 1), demonstrating the tight interactions between NCp7 and the viral RNA. UV cross-linking also generated a minor fraction of nucleoprotein complexes of about 110 kDa (lane 1). Upon prolonged incubation with T_1 RNase these 110 kDa complexes disappeared (data not shown), suggesting that NCp7 molecules were cross-linked to structured RNA sequences (see also de Rocquigny *et al.*, 1992;

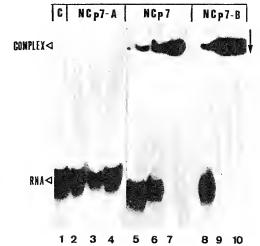
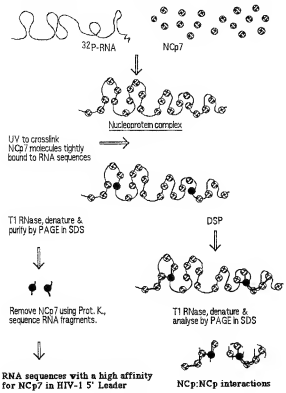
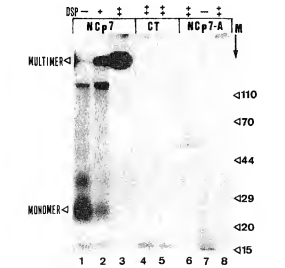


Figure 1. Gel retardation assays of HIV-1 nucleoprotein complexes formed *in vitro*. HIV-1 RNA (from positions 1 to 415, thus corresponding to the 5' leader and 5' gag sequences) labelled with [^{32}P]UMP was generated by *in vitro* transcription (see Materials and Methods). HIV-1 NCp7, NCp7-A and NCp7-B were synthesized using the Fmoc chemistry and purified by high pressure liquid chromatography (HPLC) (see Materials and Methods). Conditions of incubation and analysis are described in Materials and Methods. Lane 1, control RNA; lanes 2 to 4, NCp7-A at protein/RNA molar ratios of 6:1, 15:1 and 30:1, respectively. Lanes 5 to 7, NCp7 at protein/RNA molar ratios of 5:1, 12:1 and 25:1, respectively. Lanes 8 to 10, NCp7-B at protein/RNA molar ratios of 6:1, 15:1 and 30:1, respectively.



(a)



(b)

Figure 2. NCp7/NCp7 protein interactions revealed by the cross-linking agent DSP. (a) Scheme outlining the formation of the viral NCp7/RNA complexes and subsequent treatments with ultraviolet light (UV at 252 nm) and DSP, the homobifunctional cross-linking reagent. NCp7 molecules UV cross-linked to the viral RNA are represented by black circles. DSP-generated cross-links between NCp7 molecules are indicated by short straight lines. Viral RNA sequences with a high affinity for NCp7 are in the encapsidation-dimerization (E-DLS) element (Darlix *et al.*, 1990; Lever *et al.*, 1989). (b) ^{32}P-labelled HIV-1 RNA (1 to 415) and NCp7 were incubated under conditions

Baudin *et al.*, 1993). Treatment with DSP resulted in the formation of multimers of the RNAoligo/NCp7 complexes (lanes 2 and 3; multimer), indicating that NCp/NCp7 interactions also occur in the nucleoprotein complexes. No strong protein–protein interactions seem to take place between free NCp7 protein molecules (Morellet *et al.*, 1992). Neither RNAoligo/NC nucleoprotein monomers nor multimers were observed with the NCp7-A peptide (lanes 7 and 8).

To examine the interactions between NCp7 molecules in a cellular context, we used the yeast two-hybrid system (Fields & Song, 1989; van Aelst *et al.*, 1993); see Materials and Methods for details. Upon co-expression of the LexA-NCp7 and GAL4-NCp7 fusion proteins, yeast cells exhibited a $His^+/LacZ^+$ phenotype (Table 1), clearly indicating that NCp7/NCp7 interactions were functional *in vivo*. Expression of the mutant NCp7-A fusion proteins LexA-NCp7-A and GAL4-NCp7-A in yeast cells did not result in the growth of $His^+/LacZ^+$ cells, confirming that NCp7/NCp7 interactions were responsible for the positive $His^+/LacZ^+$ phenotype observed with the wild-type fusion proteins. Co-expression of LexA-NCp7-A and GAL4-NCp7, and *vice versa*, also did not result in the growth of $His^+/LacZ^+$ cells (Table 1).

Taken together, these results clearly indicate that functional NCp7/NCp7 interactions can take place both *in vitro* and in a cellular context. Moreover, the N-terminal domain and the first zinc finger of NCp7 appear to be of importance for these interactions (Table 1; Figure 2B, lanes 6 to 8).

Partial protection of the viral RNA against nuclease degradation in HIV-1 nucleoprotein complexes formed *in vitro*

To determine if viral RNA is protected against nuclease degradation within nucleoprotein complexes as it is in viral particles (Coffin, 1985), the complexes were treated with T_1 RNase followed by incubation with 1% SDS, extraction with phenol and analysis of the [^{32}P]RNA by PAGE in 7 M urea. Within the nucleoprotein complexes viral RNA was resistant to low or moderate concentrations of T_1 RNase (Figure 3, compare lanes controls and

reported in Materials and Methods. An NCp7/RNA molar ratio of 30:1 was used and very similar results were obtained with molar ratios of 20:1 to 40:1. Nucleocapsid complexes were irradiated with UV at 252 nm, then incubated with DSP and reactions were processed as indicated in Materials and Methods. All lanes are with [^{32}P]RNA. Lanes CT (4 and 5), controls with 100 μM DSP and UV (lane 4),and with 100 μM DSP and NCp7 (lane 5). Lanes NCp7, UV (lane 1), UV and 50 μM DSP (lane 2), and UV and DSP 100 μM (lane 3). Lanes NCp7-A, 100 μM DSP (lane 6), UV (lane 7), and UV and 100 μM DSP (lane 8). Lane M, molecular mass markers in kDa. Note that the band at about 110 kDa (lane 1) is at the interphase between the 5% and 12% gels in this PAGE analysis and represents 5 to 7% of the total [^{32}P]RNAoligo/NCp7 complexes.

Table 1. Identification of NCp7/NCp7 interactions using the yeast two-hybrid system

GAL4 / LexA	NCp7	NCp7-A	Raf
NCp7	+	-	-
NCp7-A	-	-	-
Ras	-	-	++

The yeast two-hybrid system used is described in Materials and Methods. LexA and GAL4 hybrid NC-proteins were expressed from pLex and pGAD, respectively. Transformants expressing NC protein fused to the LexA DNA-binding domain (LexA) are indicated by horizontal rows and those expressing proteins fused to the GAL4 activation domain (GAL4) are represented by vertical columns. Yeast transformants were assayed for β-galactosidase activity by a filter method for 24 hours at 30°C. Activity is indicated as + + and + for colonies with strong and moderate blue colours, respectively; and – for white colonies. Interaction between Ras and Raf proteins was used as a positive control. The same pattern of reactivity was obtained by testing the histidine auxotrophy of the transformants. Note that β-galactosidase activity as well as growth in the absence of histidine indicates an interaction between the hybrid NC-proteins.

NCp7). With a high concentration of T_1 RNase (3000 units/ml; see Figure 3 legend) this protection disappeared (lanes NCp7). In the presence of the NCp7-A peptide (lanes NCp7-A) or RT (lanes RT), viral RNA was not protected against RNase degradation, even with a low level of T_1 RNase (compare lanes: control, NCp7A and RT). The NCp7-B peptide, which lacks the zinc fingers, also protected viral RNA against T_1 RNase degradation, though this protection was consistently less effective than that by wild-type NCp7 (compare lanes NCp7-B and NCp7). These results show that viral RNA within nucleoprotein complexes is protected against nuclease degradation, although this protection is incomplete as observed in viral particles (Coffin, 1985; Darlix, 1991).

Viral DNA synthesis in HIV-1 nucleoprotein complexes

To examine reverse transcription of the viral RNA in the nucleoprotein complexes, we used, as a viral template/primer system, an RNA representing the 3' end of the genome (3' env-U3-R-poly(A)) and a DNA primer corresponding to the 3' end of the minus strand strong stop cDNA (ss-cDNA) and thus complementary to the TAR sequence. The ^{32}P-labelled primer DNA was prehybridized to the viral RNA by heat. NCp7 was then added (to half of the reactions) to form nucleoprotein complexes, and viral DNA synthesis started upon the addition of RT and dNTPs. As shown in Figure 4, viral cDNA synthesis was rapid in the nucleoprotein complexes (lanes 7 to 11; lane 7 shows 400 to 500 nt polymerized in three minutes at 37°C) as well as with free RNA (lanes 2 to 6). Interestingly, reverse transcription appeared to be more processive within the nucleoprotein complexes than with viral RNA alone, as indicated by the absence of several stops (Figure 4, compare lanes 2 to 4 with 7 to 9; small arrows and vertical bars).

Efficient and restricted DNA strand transfer in HIV-1 nucleoprotein complexes

To analyse the transfer of ss-cDNA from the 5' to the 3' end of the genomic RNA, we developed a model system (Darlix *et al.*, 1993). Two *in vitro* generated HIV-1 RNAs, representing the 5' and 3' regions of the genome, were mixed at a ratio of 1:1 with a [^{32}P]DNA primer complementary to the PBS, HIV-1 NCp7 and RTp66/p51 to form nucleoprotein complexes under physiological conditions (Darlix *et al.*, 1993; Allain *et al.*, 1994; Lapadat-Tapolsky *et al.*, 1993).

Within nucleoprotein complexes formed in the presence of only the 5' RNA, RT was able to synthesize a high level of ss-cDNA of 200 nt (Figure 5, lanes 2 and 3). In the absence of NCp7 the level

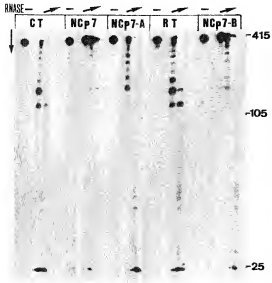


Figure 3. Protection of HIV-1 RNA within nucleoprotein complexes against RNase degradation. Nucleoprotein complexes were formed with HIV-1 [^{32}P]RNA (1 to 415) and NC protein as described in Materials and Methods. RNase (–) indicates the absence of T_1 RNase, and RNase with an arrow indicates increasing amounts of T_1 RNase. Lanes CT, controls without NCp7; lanes NCp7, with NCp7/RNA molar ratio of 30:1; lanes NCp7-A, with NCp7-A/RNA molar ratio of 40:1; lanes RT, with RT/RNA molar ratio of 20:1; lanes NCp7-B, with NCp7-B/RNA molar ratio of 35:1.

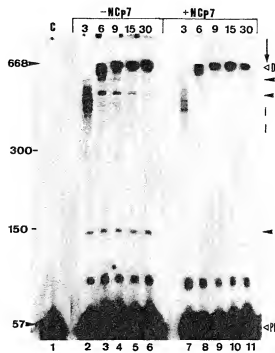


Figure 4. Reverse transcription in the HIV-1 nucleoprotein complexes formed *in vitro*. Conditions of nucleoprotein formation were as described in Materials and Methods, except that the HIV-1 RNA corresponded to the 3' region of the genome (positions 8600 to 9280). The NCp7/RNA molar ratio was 30:1. A ^{32}P-labelled DNA oligonucleotide (56 nt) complementary to the TAR sequence was used as the primer for the reverse transcription reaction. The primer/template molar ratio was 4:1 and in order to examine the reverse transcription of the viral RNA without or with NCp7 protein, the DNA primer was hybridized to the RNA by a 30 minute incubation at 65°C. Reverse transcription was in conditions described in Materials and Methods. After electrophoresis the gel was fixed, dried and autoradiographed for six hours. Size markers in nt are indicated on the left. Primer and DNA indicate the [^{32}P]DNA primer and the minus strand DNA product, respectively. Lane 1, control with RT alone and where the primer was not hybridized to viral RNA. Lanes 2 to 6 (without NCp7); and 7 to 11 (with NCp7), with the primer annealed to the RNA template and RT. The 3' end of strong stop DNA of 145 nt appears to be between positions 9040 and 9050 of the genomic RNA, in a region rich in G residues (data not shown). Note the absence of several stops with cDNA synthesis (arrows and vertical bars) in the presence of NCp7.

of ss-cDNA synthesis was about 100-fold lower (lanes 1 and 4). When both 5' and 3' RNAs were present in the nucleoprotein complexes, two minus strand DNAs were synthesized: (1) the ss-cDNA as before; and (2) the elongated minus strand DNA of 1280 nt resulting from the transfer of the ss-cDNA from the 5' RNA template to the 3' RNA and its elongation to the 5' end of the 3' RNA by RT (DNA(5'–3')). DNA strand transfer occurred within the nucleoprotein complexes (lanes 5 to 7) but not in the absence of NCp7 (compare lanes 4 and 6). The DNA strand transfer appeared to be optimal at about one NCp7 per 10 nt (Figure 5, lane 6), which represents the NCp7/RNA ratio thought to prevail in the virion nucleocapsid (Dickson *et al.*, 1985; Chen *et al.*, 1980). To confirm that no DNA strand transfer occurred in the absence of NCp7, we performed *in vitro* amplification (PCR) of the elongated minus strand DNA using two DNA oligonucleotides, one complementary to the PBS (positions 180 to 199) and the other homologous to the 5' end of the 3' RNA to synthesize a DNA of 703 base-pairs. Under our *in vitro* conditions, amplified DNA was observed only in the presence of NCp7 (data not shown).

The pioneering work of Temin and colleagues on retrovirus recombination (Hu & Temin, 1990, and references therein) demonstrated nicely that reverse transcription is restricted to the genomic RNA packaged in the virion. To investigate the possibility that reverse transcription is restricted to *in vitro* generated nucleoprotein complexes, two complexes were formed: (1) one containing the 5' RNA (5' complex); and (2) the other containing the 3' RNA (3' complex). Each nucleoprotein complex has been shown independently to support reverse transcription as indicated in Figures 4 (lanes 8 and 9) and 5 (lanes 2 and 3). The 5' and 3' complexes (formed independently) were combined at a 1:1 ratio and RT and dNTPs were added. Under these conditions, reverse transcription of the 3' RNA was dependent upon the translocation of the ss-cDNA synthesized in the 5' complex to the 3' complex. Results show

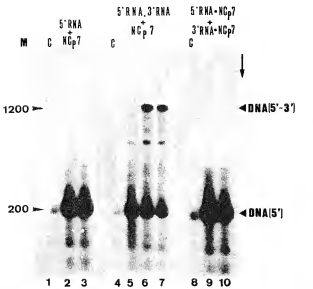


Figure 5. DNA strand transfer within nucleoprotein complexes but not between two separate complexes. HIV-1 nucleoprotein formation and conditions of reverse transcription with RTp66/p51 were as described in Materials and Methods. The primer was a ^{32}P-labelled DNA oligonucleotide of 18 nt complementary to the PBS and the 3' RNA of 1219 nt contained 3' Pol sequences (positions 3699 to 4327) in addition to 3' genomic sequences (position 8660 to poly(A)) (see Materials and Methods). Reverse transcription reactions were with a primer/RNA ratio of 4:1 and an RT to RNA ratio of 5:1 for 30 minutes at 37°C. Reaction mixtures were processed as indicated in Materials and Methods and the DNA was analysed by 1.5% agarose gel electrophoresis. The gel was first stained with ethidium bromide to identify positions of the nucleic acid markers, then fixed, dried and autoradiographed for six hours. M corresponds to the size markers in nt. All lanes contain ^{32}P-labelled primer, HIV-1 RNA and RT. Lanes 1 to 3, 5' RNA without NCp7 (lane 1) or with NCp7 at a molar ratio of 20:1 (lane 2) or 40:1 (lane 3). Lanes 4 to 7, 5' and 3' RNA without NCp7 (lane 4) or with an NCp7/RNA molar ratio of 20:1 (lane 5), 30:1 (lane 6) or 50:1 (lane 7). Lanes 8 to 10, 5' RNA and 3' RNA were preincubated independently with NCp7 (two different tubes) for five minutes at 37°C. The reactions were then combined, and RT and dNTPs added; without NCp7 (lane 8) or with NCp7 at a molar ratio of 30:1 (lane 9) or 50:1 (lane 10). DNA(5') and DNA(5'-3') correspond to ss-cDNA (200 nt) and transferred and elongated minus strand DNA (1280 nt), respectively. Note that although ss-cDNA synthesis occurred, no elongated minus strand DNA was synthesized (lanes 9 & 10). Amplification by PCR of the HIV-1 DNA products: PCR was for 15 cycles under conditions described by Darlix *et al.* (1993) using 2% of the DNA products and two DNA oligonucleotides, one complementary to the PBS (GTCCCTGTTCGGGCGCCA) and the other homologous to a 3' Pol sequence of the 3' RNA (GACAGCAGTACAAATGGC).

that ss-cDNA synthesis occurred in the 5' complex (Figure 5, lanes 9 and 10) but DNA strand transfer and subsequent extension were undetectable under these conditions (compare lanes 9 to 10 with 6 to 7). As described above, the viral DNA synthesized was subjected to amplification by PCR (see legend to Figure 5) to determine if low levels of DNA strand transfer (undetectable by PAGE–urea analysis) occurred. Results indicate that DNA strand transfer between the 5' and the 3' nucleoprotein complexes was reduced to about 1 to 2% of that within a complex containing both RNAs (data not shown). DNA strand transfer from a 5' to free 3' RNA was also examined by PCR amplification and found to be very low (data not shown).

Discussion

The nucleocapsid of HIV-1, like that of the avian and murine oncoretroviruses, contains two major structural components, the nucleocapsid protein and the dimeric RNA genome, as well as reverse transcriptase and integrase. The HIV-1 nucleocapsid substructure appears to result from the tight association of about 1500 to 2000 molecules of NCp7 and one dimeric RNA (Chen *et al.*, 1980; Coffin, 1985; Darlix *et al.*, 1990, 1991). The numerous interactions between the NCp7 molecules and the genomic RNA within this specialized nucleoprotein structure are poorly understood. The possible interactions between NCp7 and the RT enzyme (Barat *et al.*, 1993), and their co-operation during the reverse transcription process (Barat *et al.*, 1989; Darlix *et al.*, 1993; Allain *et al.*, 1994) are also poorly understood. In order to investigate the molecular interactions within the retroviral nucleocapsid, and how these interactions control the stability of this nucleoprotein structure as well as the reverse transcription and recombination processes, attempts were made to reconstitute functional retroviral nucleocapsids *in vitro*.

Results presented here reveal that stable and functional nucleoprotein complexes resembling the HIV-1 nucleocapsid have been formed. The stability of the *in vitro* generated nucleoprotein complexes appears to be achieved through the binding of a large number of NCp7 molecules to the viral RNA (Figure 1) and NCp7/NCp7 interactions (Figure 2 and Table 1). While NCp7 seems to be mostly monomeric in solution (Morellet *et al.*, 1992), protein–protein interactions take place upon binding of NCp7 to the RNA (Figure 2). Such a multimeric organization of NCp7 molecules on and along the viral RNA most likely accounts for the protection of the viral RNA against RNase degradation (Figure 3). Despite the fact that the viral RNA is covered with NCp7 molecules and that small enzymes such as T_1 RNase have limited access to RNA within the nucleoprotein complexes, RT, a 117 kDa (p66/p51) protein, was able efficiently to initiate reverse transcription and elongate cDNA by copying the viral RNA template (Figure 4). In addition, reverse transcription appeared to be more processive when cDNA synthesis occurred within the nucleoprotein complexes (Figure 4). This improved processivity is probably promoted by the RNA unwinding activity of NCp7 (Khan & Giedroc, 1992). Also, postulated RT/NC protein interactions (Barat *et al.*, 1993) may explain the facility with which RT interacts with viral

RNA within nucleoprotein complexes; this is under investigation.

Synthesis of a complete proviral DNA necessitates strand transfers, the first of which corresponds to the minus strand DNA transfer from the 5' to the 3' end of the RNA genome (Varmus & Swanstrom, 1985; Coffin, 1985; Darlix *et al.*, 1993). A high level of DNA strand transfer from a 5' viral RNA template to a 3' viral RNA and corresponding to the first DNA strand transfer was achieved in the nucleoprotein complexes generated *in vitro* (Figure 5). However, DNA strand transfer was drastically reduced when the 5' and 3' viral RNAs were in separate nucleoprotein complexes (Figure 5). This resembles the situation *in vivo*, where DNA strand transfers resulting in genetic recombination occur between two different retroviral genomes only when virus particles contain two different RNAs in the form of heterodimers. These heterozygous viruses can be formed when the parental cell harbours two genetically distinct proviruses (Hu & Temin, 1990, and references therein). However, recombination between two genetically different retroviral genomes does not take place when homozygous viruses containing genetically different RNAs, in the form of homodimers, are mixed and used to infect fresh cells (Hu & Temin, 1990).

Work is in progress to define more precisely the interactions between NCp7 protein molecules in the nucleocapsid substructure and to characterize agents capable of inhibiting these interactions.

Materials and Methods

Plasmid DNA construction and preparation

Standard procedures were used for restriction nuclease digestion and plasmid DNA construction (Sambrook *et al.*, 1989). *Escherichia coli* HB-101 strain 1035 was used for plasmid DNA amplification.

Plasmid construction for the synthesis of large amounts of HIV1 5' RNA and 3' RNA has been reported. pHIVCG-3 was used for 5' RNA (Darlix *et al.*, 1990) and pHIVCG-8.6 for 3' RNA (Darlix *et al.*, 1993). In addition, pHIVCG-46 was constructed for the *in vitro* synthesis of 3' RNA with pol sequences. This was done by inserting pol sequences (*Kpn*I and *Dra*I sites at positions 3699 to 4327, respectively) into pHIVCG-8.6 cut with *Kpn*I and *Dra*I. The resulting construct pHIVCG-46 contained pol sequences inserted at the 5' end (position 8660) of pHIVCG-8.6.

RNAs generated *in vitro*

5' HIV1 RNA was generated *in vitro* using T7 RNA polymerase and pHIVCG-3 DNA and corresponded to positions 1 to 415 (*Hae*III site) (R, U5, PBS, E-DLS region and 5' gag). 3' HIV-1 RNAs from positions 8600 to 9229 (3' env, ppt, U3, R, a 20 nt poly(A) tail and additional nt to insert in the *Xho*I site) and 3699 to 4327 to 8660 to 9229 (3' pol, 3' env, ppt, U3, R, poly(A) and additional nt to insert in the *Xho*I site) were synthesized by T3 RNA polymerase using pHIVCG-8.6 and pHIVCG-46 cut with *Xho*I, respectively. 5' and 3' HIV-1 RNAs were synthesized under conditions described by Darlix *et al.* (1993), purified by two extractions with phenol and two precipitations with ethanol, and dissolved in sterile water at 1 mg/ml. The RNAs were analysed by 1% agarose gel electrophoresis in 50 mM Tris-borate (pH 8.3). Replication primer $tRNA^{Lys,3}$ labelled with [^{32}P]UTP was generated using the pTL9 construct as described (Barat *et al.*, 1991). [^{32}P]-$tRNA^{Lys,3}$ was purified by 10% polyacrylamide gel electrophoresis containing 7 M urea, and subsequently heat denatured and cooled down slowly before use. DNA oligonucleotide primers, complementary to either the PBS (18 nt) or the 3' TAR sequence (66 nt), were also used for the reverse transcription reactions. The DNA primers were (5'-^{32}P)-labelled using [γ-^{32}P]ATP and bacteriophage T4 kinase, and purified by 12% PAGE in 7 M urea.

NCp7 protein, NCp7 mutants and RTp66/p51

NCp7 (72 amino acid residues) and NCp7 mutants were prepared chemically by peptide synthesis as described (de Rocquigny *et al.*, 1992). The NCp7 mutants used were: NCp7A in which the first 28 residues had been deleted and NCp7B (47 amino acids) in which both zinc fingers had been removed.

The heterodimeric enzyme reverse transcriptase RTp66-p51, a gift from S. Le Grice, was produced in *E. coli* and purified (Barat *et al.*, 1989).

T_1 RNase was from Sigma and dithiobis(succinimidyl propionate) was from Pierce.

HIV-1 nucleoprotein formation

HIV-1 RNA (from positions 1 to 415) labelled with [^{32}P]UMP (50 nM) was incubated without or with NCp7, NCp7A or NCp7B in 20 µl of 30 mM Tris-HCl (pH 7.1), 30 mM NaCl, 0.2 mM $MgCl_2$, 10 µM $ZnCl_2$, 5 mM dithiothreitol and 10 units of RNAsin (Promega) for five minutes at 37°C. Then, SDS (0.1% final) and glycerol (5% final) were added and electrophoresis was carried out in 6% polyacrylamide gels in 0.1% SDS. After electrophoresis, the gel was exposed for two hours at 4°C.

NCp7/NCp7 protein interactions

^{32}P-labelled HIV-1 RNA (1 to 415) and NCp7 were incubated under exactly the same conditions as described above, at an NCp7 to RNA molar ratio of 20:1 to 40:1. Nucleocapsid complexes were irradiated with UV at 252 nm (20 W lamp, at 2 cm) for 40 seconds at 20°C and then incubated with DSP, a protein cross-linking agent (Schweizer *et al.*, 1982), at 50 or 100 µM for 30 minutes at 20°C. Reactions were stopped by the addition of Tris-HCl (pH 7.0; 0.3 M, final concn) for 15 minutes and heated at 95°C for two minutes. Assays were done at 4°C, 30 units of T_1 RNase was added and the reactions were incubated for 30 minutes at 4°C, then 30 minutes at 20°C. Next, glycerol, SDS and 2-mercaptoethanol were added at final concentrations of 5%, 2% and 0.1%, respectively, and the reactions were heated at 55°C for five minutes. Analysis was performed by 12% PAGE in 0.2% SDS. After electrophoresis, the gel was fixed, dried and autoradiographed for three hours.

NCp7/NCp7 interactions analysed by means of the yeast two-hybrid system

Plasmids pLex-Ras and pGAD-Raf, expressing extraneous targets (Fields & Song, 1989; van Aelst *et al.*, 1993), as well as pLex-10 and pGAD, were provided by R. Benarous (Paris). pLex-10 derives from pBTM116 (Bartel

et al., 1993) and pGAD-1318 from pGAD.GH (Breeden & Nasmyth, 1985). These derived vectors present the *Bam*HI site in a different reading frame. To construct pLex-NC and pGAD-NC, appropriate synthetic oligonucleotide primers with a *Bam*HI site were used to amplify the NCp7 coding sequence by PCR. DNA products were treated with Klenow fragment and digested with *Bam*HI, resulting in a fragment that had one blunt end. This was ligated to pLex-10 or pGAD-1318, which had been digested with *Pst*I or *Xho*I, respectively, treated with Klenow fragment and digested with *Bam*HI. The yeast reporter strain L40 (Vojtek *et al.*, 1993) containing two inducible genes, *HIS3* and *LacZ*, was transformed simultaneously with the plasmids containing the hybrid NC sequence. After three days of growth on drop-out medium lacking tryptophan and leucine, double transformants were replica-plated either on selective medium lacking histidine or on Whatman 40 filters, and tested for β-galactosidase activity (Breeden & Nasmyth, 1985).

Reverse transcription and analysis of the cDNA products

5' and 3' RNAs generated *in vitro* were mixed at a ratio of 1:1 and heat denatured for one minute at 80°C. Primer DNA (5'-^{32}P)-labelled, complementary to PBS or TAR, was annealed to the template by a 30 minute incubation at 65°C. Hybridization of the natural primer $tRNA^{Lys,3}$ to the PBS was conducted with NCp7 protein. 5' and 3' RNAs (60 nM each) were incubated with primer (100 nM) and NCp7 (0.5 to 2 μM) in 10 μl final volume for 15 minutes at 37°C. Subsequently, the reaction was diluted to 25 μl in 30 mM Tris-HCl (pH 8.0), 50 mM NaCl, 5 mM dithiothreitol, 3 mM $MgCl_2$, the four deoxyribonucleotides were added at 0.25 mM and RTp66/p51 at 80 nM, and the reaction was incubated for an additional 30 minutes. Reactions were stopped with 10 mM EDTA, 1% SDS, heated for two minutes at 50°C, and extracted twice with phenol and phenol/chloroform (1:1, v/v). The supernatant phase was recovered and precipitated with three volumes of ethanol. The ^{32}P-labelled DNA recovered by centrifugation was dissolved in 80% formamide, heat denatured for one minute at 98°C and analysed by agarose or polyacrylamide gel electrophoresis. At the end of the electrophoresis the gel was stained with ethidium bromide (0.5 μg/ml for five minutes) and visualized, and subsequently fixed with 5% trichloroacetic acid, dried and autoradiographed.

Digestion of HIV-1 RNA present in nucleoprotein complexes by T_1 RNase

Nucleoprotein complexes were formed with HIV-1 [^{32}P]RNA (1 to 415) as described above and then incubated in the presence of 0.3, 3 and 30 units T_1 RNase (Sigma) per reaction for 30 minutes at 20°C. Then, 2% SDS was added to the reactions and RNA was extracted with phenol, then phenol-chloroform (1:1, v/v) and analysed by 6% PAGE in 7 M urea. Autoradiography of the gel was for six hours.

Acknowledgements

Thanks are due to Bernard Roques and Hugues de Rocquigny (Paris) for NCp7, NCp7-A and NCp7-B; Suart Le Grice (Case Western, USA) for HIV-1 reverse transcriptase; Richard Benarous and Serge Benichou (Paris) for support with the yeast two hybrid system; and Mary Lapadat-Tapolsky (Lyon) for a critical reading of the manuscript. This work was supported by grants from the French program against AIDS (ANRS), MGEN, Rhône-Poulenc Rorer and ECS.

References

Allain, B., Lapadat-Tapolsky, M., Berlioz, C. & Darlix, J.-L. (1994). Transactivation of the minus-strand DNA transfer by nucleocapsid protein during reverse transcription of the retroviral genome. *EMBO J.* 13, 973–981.

Barat, C., Lullien, V., Schatz, O., Grüninger-Leitch, M., Le Grice, S., Nugeyre, M.-T., Barré-Sinoussi, F. & Darlix, J.-L. (1989). HIV-1 reverse transcriptase specifically interacts with the anti-codon domain of its cognate primer tRNA. *EMBO J.* 8, 3279–3285.

Barat, C., Le Grice, S. & Darlix, J.-L. (1991). Interaction of HIV-1 reverse transcriptase with a synthetic form of its replication primer $tRNA^{Lys,3}$. *Nucl. Acids Res.* 19, 751–757.

Barat, C., Schatz, O., Le Grice, S. & Darlix, J.-L. (1993). Analysis of the interactions of HIV1 replication primer $tRNA^{Lys,3}$ with nucleocapsid protein and reverse transcriptase. *J. Mol. Biol.* 231, 185–190.

Bartel, P. L., Chien, C. T., Sternglanz, R. & Fields, S. (1993). *Cellular Interactions in Development: A Practical Approach* (Hartley D. A., ed.), p. 153, Oxford University Press, Oxford.

Baudin, F., Marquet, R., Isel, C., Darlix, J.-L., Ehresmann, B. & Ehresmann, Ch. (1993). Functional sites in the 5' region of HIV1 RNA form defined structural domains. *J. Mol. Biol.* 229, 382–397.

Bieth, E., Gabus, C. & Darlix, J.-L. (1990). A study of the dimer formation of Rous sarcoma virus RNA and of its effect on viral protein synthesis *in vitro*. *Nucl. Acids Res.* 18, 119–126.

Breeden, L. & Nasmyth, K. (1985). Regulation of the yeast HO gene. *Cold Spring Harbor Symp. Quant. Biol.* 50, 643–650.

Chen, M., Garon, C. & Papas, T. (1980). Native ribonucleoprotein is an efficient transcriptional complex of avian myeloblastosis virus. *Proc. Natl Acad. Sci. USA*, 77 1296–1300.

Coffin, J. M. (1979). Structural replication and recombination of retrovirus genomes: some unifying hypotheses. *J. Gen. Virol.* 42, 1–26.

Coffin, J. (1985). Genome structure. In *RNA Tumor Viruses* (Weiss, R., Teich, N., Varmus, H. & Coffin, J., eds), part 2, 2nd edit., pp. 17–74. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Dannull, J., Surovoy, A., Jung, G. & Moelling, K. (1994). Specific binding of HIV-1 nucleocapsid protein to PSI RNA *in vitro* requires N-terminal zinc finger and flanking basic amino acid residues. *EMBO J.* 13, 1525–1533.

Darlix, J.-L. (1991). Eléments de structure et de variabilité du rétrovirus humain HIV-1 *Bull. Inst. Pasteur*, 89, 211–242.

Darlix, J.-L. & Spahr, P.-F. (1982). Binding sites of viral protein p19 onto RSV RNA and possible control of viral functions. *J. Mol. Biol.* 160, 147–161.

Darlix, J.-L., Gabus, C., Nugeyre, M.-T., Clavel, F. & Barré-Sinoussi, F. (1990). *Cis* elements and *trans* acting factors involved in the RNA dimerization of HIV-1. *J. Mol. Biol.* 216, 689–699.

Darlix, J.-L., Vincent, A., Gabus, C., de Rocquigny, H. & Roques, B. (1993). *Trans*-activation of the 5' to 3' viral DNA strand transfer by nucleocapsid protein during

reverse transcription of HIV1 RNA. *C. R. Acad. Sci. Life Sci.* **316**, 763–771.

de Rocquigny, H., Gabus, C., Vincent, A., Fournié-Zaluski, M.-C., Roques, B. & Darlix, J.-L. (1992). Viral RNA annealing activities of HIV-1 nucleocapsid protein require only peptide domains outside the zinc fingers. *Proc. Natl Acad. Sci. USA*, **89** 6472–6476.

de Rocquigny, H., Ficheux, D., Gabus, C., Allain, B., Fournié-Zaluski, M.-C., Darlix, J.-L. & Roques, B. (1993). Two short basic sequences surrounding the zinc finger of NCp10 protein of MoMuLV are critical for RNA annealing activity. *Nucl. Acids Res.* **21**, 823–829.

Dib-Hajj, F., Khan, R. & Giedroc, D. P. (1993). Retroviral nucleocapsid proteins possess potent nucleic acid strand renaturation activity. *Protein Sci.* **2**. 231–243.

Dickson, C., Elsenman, R., Fan, H., Hunter, E. & Teich, N. (1985). Protein biosynthesis and assembly. In *RNA Tumor Viruses* (Weiss, R., Teich, N., Varmus. H. & Coffin, J., eds), part 2, 2nd edit., pp. 513–648, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Fields, S. & Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature*, **340**, 245–246.

Hu, W. & Temin, H. (1990). Genetic consequence of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. *Proc. Natl Acad. Sci. USA*, **87**, 1556–1560.

Khan, R. & Giedroc, D. (1992). Recombinant HIV-1 nucleocapsid (NCp7) protein unwinds tRNA. *J. Biol. Chem.* **267**, 6689–6695.

Lapadat-Tapolsky, M., de Rocquigny, H., van Gent, D., Roques, B., Plasterk, R. & Darlix, J.-L. (1993). Interactions between HIV-1 nucleocapsid protein and viral DNA may have important functions in the viral life cycle. *Nucl. Acids Res.* **21**, 831–839.

Lever, A., Göttlinger, H. G., Haseltine, W. & Sodroski, J. (1989). Identification of a sequence required for efficient packaging of HIV-1 RNA into virions. *J. Virol.* **63**, 4085–4087.

Méric, C., Darlix, J. L. & Spahr, P.-F. (1984). It is Rous sarcoma virus protein p12 and not p19 that binds tightly to Rous sarcoma virus RNA. *J. Mol. Biol.* **173**, 531–538.

Méric, C., Gouilloud, E. & Spahr, P.-F. (1988). Mutations in Rous sarcoma virus nucleocapsid protein p12 (NC): Deletions of Cys-His boxes. *J. Virol.* **62**, 3328–3333.

Morellet, N., Jullian, N., De Rocquigny, H., Maigret, B., Darlix, J.-L. & Roques, B. (1992). ^{1}H NMR derived conformation of the nucleocapsid protein NCp7 of human immunodeficiency virus type 1. *EMBO J.* **11**. 3059–3065.

Prats, A. C., Sarih, L., Gabus, C., Litvak, S., Keith, G. & Darlix, J.-L. (1988). Small finger protein of avian and murine retroviruses has nucleic acid annealing activity and positions the replication primer tRNA onto genomic RNA. *EMBO J.* **7**, 1777–1783.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schweizer, E., Angst, W. & Lutz, H. (1982). Glycoprotein topology on intact human red blood cells reevaluated by crosslinking following amino acid group supplementation. *Biochemistry*, **21**, 6807–6818.

van Aelst, L. V., Barr, M., Marcus, S., Polverino, A. & Wigler, M. (1993). Complex formation between Ras and Raf and other protein kinases. *Proc. Natl Acad. Sci. USA*, **90**, 6213–6217.

Varmus, H. & Swanstrom, R. (1985). Replication of retroviruses. In *RNA Tumor Viruses* (Weiss, R., Teich, N., Varmus, H. & Coffin, J., eds), part 2, 2nd edit., pp. 369–512, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Votjek, A. B., Hollenberg, S. M. & Cooper, J. A. (1993). Mammalian Ras interacts directly with the serine/threonine kinase Raf. *Cell*, **74**, 205–214.

You, J. C. & McHenry, C. S. (1994). Human immunodeficiency virus nucleocapsid protein accelerates strand transfer of the terminally redundant sequences involved in reverse transcription. *J. Biol. Chem.* **269**, 31491–31495.

Edited by J. Karn

(Received 7 June 1995; accepted 19 July 1995)